

Agentes y sistemas multiagente

Modelos teóricos y aplicaciones

Marcelo Errecalde Guillermo Aguirre Alfredo Muchut Diego Devia

LIDIC *

Departamento de Informática

Universidad Nacional de San Luis (UNSL)

Ejército de los Andes 950 - Local 106

5700 - San Luis. Argentina

e-mail: {merreca, gaguirre, amuchut, dhdevia}@unsl.edu.ar

Tel. (02652)-424938. Int. 259

Resumen

Este artículo describe en forma resumida los trabajos de investigación y desarrollo que se están llevando a cabo en la línea de *Agentes y Sistemas Multiagente* del Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC) de la Universidad Nacional de San Luis. Los principales enfoques abordados en esta línea de investigación fueron presentados en el WICC2003 [5]. Se continúa ahora con el estado de avance y los resultados obtenidos recientemente con la utilización de estos enfoques. Si bien este artículo constituye una breve síntesis de los trabajos realizados y en ejecución, intentará reflejar las principales temáticas abordadas para posibilitar un intercambio de experiencias con otros investigadores participantes del Workshop, que trabajen sobre líneas de investigación afines. El artículo está organizado alrededor de dos grandes ítems: *modelos teóricos y aplicaciones*. En el primer caso, se describen los estudios y desarrollos realizados con modelos teóricos de agentes y sistemas multiagente. En las aplicaciones se presentan las principales conclusiones de las experiencias logradas con el diseño e implementación de sistemas basados en estos paradigmas.

Palabras Claves: Agentes Inteligentes, Sistemas Multiagente, Aprendizaje de Máquina.

1. Introducción

Nuestra línea de investigación está abocada a la investigación y desarrollo de sistemas estructurados en base al paradigma de agentes y sistemas multiagente. Los enfoques para la construcción de este tipo de sistemas son muy variados, demostrando cada uno de ellos una serie de falencias y ventajas dependiendo del dominio de aplicación. Por este motivo, los trabajos realizados en el grupo de investigación están orientados al estudio y aplicación de distintos paradigmas para la construcción de agentes y a su posible integración en aquellos casos que sea conveniente.

La investigación en el área de agentes y sistemas multiagente debe considerar no sólo el desarrollo de sistemas y aplicaciones concretas que resuelvan problemas particulares, sino también

*El laboratorio es subvencionado por la UNSL y la ANPCYT.

el estudio y desarrollo de modelos teóricos que puedan fundamentar y justificar la aplicación de determinadas técnicas en el desarrollo de aplicaciones prácticas. Por este motivo, los trabajos desarrollados dentro del grupo han intentado considerar ambos aspectos y la organización del artículo refleja esta visión. En primer lugar, en la sección 2 se describe un modelo de aprendizaje basado en los principios de los sistemas multiagente y un modelo de colaboración deliberativo que recientemente se ha comenzado a investigar dentro de la línea. En la sección 3 en cambio, se resumen algunas de las principales conclusiones obtenidas en el desarrollo de sistemas multiagente que están siendo utilizados actualmente por el grupo de investigación.

2. Modelos de Aprendizaje y Colaboración

2.1. Aprendizaje Basado en múltiples fuentes de experiencia

El *Aprendizaje Basado en Múltiples Fuentes de Experiencia* (ABMFE) es un modelo de aprendizaje para agentes que aprenden directamente de su interacción con el ambiente y con otros agentes. Este modelo de aprendizaje intenta dar respuesta a problemas observados en el área del *Aprendizaje por Refuerzo* (AR) [10] y el AR en sistemas multiagentes (ARMA) [1, 6]. Entre estos problemas podemos citar la necesidad de considerar el uso intensivo de información adicional variada, la ocurrencia de múltiples procesos de aprendizaje simultáneos, decisiones de alto nivel y meta-aprendizaje y el uso de mecanismos de eliminación o inhibición de alternativas y procesos.

En el contexto del ABMFE, una *fente de experiencia* es cualquier entidad o proceso que aporte información para decidir que acción tomar en las distintas situaciones o contribuya en el aprendizaje de esta tarea. *Ejemplos*: Sistema de AR, supervisado, por consejos, planificador, modelos del ambiente y agentes, leyes sociales, etc.

La idea es mantener separadas las distintas componentes (internas o externas) que influyen en el control y aprendizaje del agente. El enfoque propuesto para la modularización del agente se ubica en el contexto de los *sistemas multiagente*. Las componentes fundamentales del aprendiz son las fuentes de experiencia. Cada fuente de experiencia está representada dentro del aprendiz como un agente que puede participar de los distintos procesos de control y aprendizaje adoptando distintos roles. La sociedad formada por las distintas fuentes de experiencia estará subordinada al aprendiz como un todo. De esta manera, el aprendiz puede ser considerado como un agente abstracto construido sobre otros agentes menos abstractos.

El aprendizaje de un agente es considerado como un fenómeno social en el que pueden participar diversos agentes de la sociedad en que habita. El aprendiz forma parte de un contexto social donde puede aprender no sólo desde las recompensas recibidas por sus acciones, sino también desde las múltiples fuentes de información que puedan estar socialmente disponibles, incluyendo agentes humanos y artificiales. La perspectiva multiagente del aprendiz se aplica “hacia afuera” considerándolo parte de un sistema multiagente que lo contiene y también “hacia adentro” al considerar que sus componentes (fuentes de experiencia) constituyen a su vez un sistema multiagente.

En base al modelo de ABMFE se ha definido una arquitectura genérica que ha permitido integrar dentro de un único marco de trabajo distintas áreas vinculadas al Aprendizaje por Refuerzo que han sido consideradas usualmente en forma aislada: a) Aprendizaje por Refuerzo libre de modelo y basado en modelo, b) Programación Dinámica de Tiempo Real (RTDP), c) Aprendizaje en Juegos estocásticos, d) Exploración dirigida, no dirigida e informada, e) Exploración local y no local, f) Políticas *off-policy* y *on-policy*.

Distintos métodos existentes del AR pudieron ser integrados al contexto del ABMFE y se definieron nuevos métodos de aprendizaje como por ejemplo: a) combinación de exploración dirigida y no dirigida, b) exploración dirigida por el objetivo, c) exploración de Boltzmann informada,

d) exploración restringida, e) integración de leyes sociales y planes en el AR, f) RTDP adaptativo (on policy), g) AR basado en modelo (on policy) y h) RTDP para juegos estocásticos.

2.2. Modelo para el trabajo colaborativo

Buena parte del trabajo sobre resolución cooperativa de problemas considera una actitud benevolente: los agentes en el sistema comparten implícitamente una meta común por lo que no existen conflictos potenciales sobre los objetivos a alcanzar. Esta suposición implica que los agentes estarán dispuestos a brindar ayuda para alcanzar un objetivo común para todo el sistema, en lugar de perseguir sus metas individuales. De este modo, se simplifica mucho el diseño de los agentes [9, 7]. En contraste, si se consideran sistemas multiagente (SMA) generales, no se puede asumir que los agentes comparten un objetivo común, ya que pueden ser creados por personas independientes donde cada agente persigue sus metas individuales. Por este motivo, la investigación en estos SMA's se aboca principalmente al diseño de sociedades de agentes autónomos que puedan reconocer y resolver conflictos, con capacidad de negociar o de llegar a acuerdos [11, 8].

Si los agentes son realmente autónomos deben ser capaces de coordinar sus actividades y cooperar con otros dinámicamente. Este comportamiento dinámico requiere que los agentes se comuniquen y alcancen acuerdos satisfactorios que permitan resolver problemas en conjunto [4, 3]. La interacción entre los miembros, ya sea como conversaciones explícitas o como observación del comportamiento de los demás miembros, lleva a un intercambio de información que se traduce en alcanzar estados mentales dentro de los cuales está el conocimiento sobre el grupo. Este conocimiento abarca tanto lo que cada uno sabe individualmente como así también la existencia del grupo y la posibilidad de trabajar en equipo.

Dentro de la línea de investigación, se ha comenzado a trabajar recientemente en una arquitectura para agentes que soporte este tipo de interacciones, a partir del trabajo colaborativo de un grupo de agentes autónomos. El aspecto central para el desarrollo de la arquitectura de estos agentes será establecer la forma en que el agente internamente realiza la deliberación que define el curso de acción [2].

El aspecto reactivo del agente en relación a los cambios del entorno si bien es considerado, no es troncal. En cambio, se asignará una importancia significativa a la administración y regulación de los recursos disponibles para conseguir sus objetivos, eligiendo cuidadosamente las tareas a realizar. Esta elección normalmente se estructura en planes que sirven para dirigir el razonamiento sobre medios-y-fines, concentrándose en las opciones que sean más relevantes y manteniendo un reducido grupo de creencias que conforman la base sobre la que se realiza el razonamiento práctico. Los planes son necesarios para garantizar un cierto comportamiento estable, pero también es indispensable una frecuente reconsideración de los mismos. Actualmente, se está estudiando como una alternativa para considerar los cambios en los planes, la utilización de un sistema de filtrado basado en un razonamiento rebatible que de alguna manera permita acotar el conocimiento a ser considerado cuando se deben tomar decisiones. Las características no-monótonas de este tipo de razonamiento sirven como garante de contar con un núcleo reducido de conocimiento.

3. Aplicaciones

3.1. SMALL

Esta aplicación fue pensada para facilitar la administración de referencias a sitios de internet desde dos enfoques. Un mismo usuario en diferentes estaciones de trabajo y un grupo de usuarios que desean compartir un conjunto de sitios favoritos. Debido a estos dos enfoques fue necesario

establecer una organización del SMA que previera la posibilidad de dos formas diferentes de utilización del sistema.

La primer suborganización se establece alrededor de un agente principal destacado: el coordinador. A través de este coordinador se establecen todas las comunicaciones y se sincronizan las actividades de los llamados agentes homogéneos dentro del SMA. La perspectiva en este caso es la denominada *centrada en el espacio* [12], cada agente homogéneo posee un conjunto de funciones requeridas para toda la aplicación, pero limitada a un espacio reducido del ambiente (una estación de trabajo) donde el agente actúa. Además se mantiene una relación de subordinación estática entre el coordinador y los homogéneos ya que estos últimos siempre responderán los requerimientos del primero.

La otra suborganización se puede describir como una arquitectura tipo blackboard. Un agente administrador es responsable de las modificaciones sobre los datos compartidos. Los demás agentes representan las fuentes de conocimiento que proveen información sobre el ambiente, logrando que el resto del sistema perciba el entorno. Con este agente administrador se garantizó la consistencia de la información compartida y el secuenciamiento de los accesos.

La organización del SMA incluye dos suborganizaciones que conviven mientras el sistema está en uso. Los agentes tienen dos roles e intervienen en diferentes conversaciones según se den las circunstancias.

3.2. AULERO

A partir del análisis realizado sobre la problemática de la administración de aulas, se determinó que la situación a resolver en sí misma es distribuida. Las cátedras y auleros pueden ser representados mediante agentes individuales. Ese mismo carácter distribuido también está presente en la información que cada uno de los agentes debe administrar. Es por esta característica distribuida, que la organización del SMA incluye dentro de los agentes cátedra toda una técnica de votación para resolver los conflictos entre las preferencias establecidas. Esta técnica mostró un comportamiento adecuado, permitiendo que cada agente cátedra determinara por sí mismo la elección más conveniente entre diversas alternativas. De este modo se asegura satisfacer las expectativas del usuario mediante la conciliación de preferencias.

Por otro lado, ciertas actividades requieren que los agentes coordinen sus acciones para satisfacer determinadas metas globales. Así fue necesario establecer ciertos *compromisos* y *convenios* para superar dependencias en las acciones, encontrar patrones globales y debido a que ningún agente por sí mismo posee suficientes capacidades, recursos o información para alcanzar las metas del sistema. Los compromisos son promesas de realizar una tarea de una forma determinada y las convenciones establecen la manera de administrar los compromisos en circunstancias cambiantes. Estos convenios y compromisos permitieron un adecuado funcionamiento del sistema.

3.3. Control y aprendizaje de robots móviles

Tanto los sistemas basados en lógica difusa como en redes neuronales han sido aplicados exitosamente a problemas de control. Cada uno presenta ventajas y desventajas. Dentro de la línea de investigación se comparó el desempeño de soluciones basadas en redes neuronales convencionales con las obtenidas usando sistemas basados en lógica difusa para problemas de control.

El problema de control para el cual se desarrollaron controladores de distinta índole (difusos y neuronales) consistieron básicamente en problemas donde el robot debe aproximarse a fuentes de luz en un mundo sin obstáculos (ambiente estático) o debe apagar todas las luces insertas en el mundo (ambiente dinámico).

El controlador difuso fue desarrollado en base a la propuesta de Mamdani. También se probó el método SAM y la incorporación de reglas de contexto. Las salidas del controlador difuso fueron utilizadas para entrenar distintos tipos de redes neuronales. El objetivo de esta tarea fue el de analizar en que medida la red neuronal podía aproximar el comportamiento del controlador difuso, a medida que a este último se le incorporaban características más complejas.

Se comprobó que el método de defusificación *Centro de Gravedad* produce un comportamiento más continuo o natural que los obtenidos con *Máximo* y *Media de Máximos*, si bien su costo computacional es más alto. La utilización de reglas de contexto fue una solución práctica a problemas que una regla difusa común no podía resolver, por ejemplo, cuando el robot estaba en oscuridad y se necesitaba un comportamiento exploratorio para encontrar las luces y un comportamiento antiestancamiento para no trabarse contra las paredes.

Se logró obtener redes con un comportamiento similar al del controlador difuso incluso en aquellos casos en que se incorporaron características complejas como por ejemplo las reglas de contexto. Un detalle importante en este sentido es que las redes entrenadas mostraban una velocidad superior para generar la salida que la mostrada por el controlador difuso. Estos resultados nos llevan a considerar como una alternativa para el diseño del controlador para un robot, el entrenamiento de redes neuronales con el comportamiento del controlador difuso en el cual es relativamente directo incorporar el conocimiento de un experto; estas redes, no sólo tendrán tiempos de respuestas mejores que los del controlador difuso sino que además podrán seguir aprendiendo en su interacción con el ambiente.

Referencias

- [1] Michael Bowling and Manuela Veloso. An analysis of stochastic game theory for multiagent reinforcement learning. Technical report, Computer Science Department, Carnegie Mellon University, 2000.
- [2] Michael Btarmen, David Isreal, and Martha E. Pollack. Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4(4):349–355, 1988.
- [3] Philip R. Cohen and Hector J. Levesque. Communicative actions for artificial agents. In Victor Lesser and Les Gasser, editors, *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS'95)*, pages 65–72, San Francisco, CA, USA, 1995. The MIT Press: Cambridge, MA, USA.
- [4] Philip R. Cohen, Jerry Morgan, and Martha E. Pollack. *Intentions in Communication*. The MIT Press: Cambridge, MA, USA, Cambridge, MA, USA, 1988.
- [5] Marcelo Errecalde, Guillermo Aguirre, Fernando Gonzalez, Diego Devia, and Cristina Saez. Un resumen de experiencias y el estado de avance en el uso de agentes y sistemas multiagente. In *Anales del V Workshop de Investigadores en Ciencias de la Computación 2003 - WICC 2003*, pages 215–219, Tandil, Argentina, 2003.
- [6] Junling Hu. *Learning in Dynamic Noncooperative Multiagent Systems*. PhD thesis, Computer Science and Engineering, The University of Michigan, 1999.
- [7] Sarit Kraus. Interaction and collaboration in multi-agent systems. In A. G. Cohn, editor, *Proceedings of the Eleventh European Conference on Artificial Intelligence*, pages 823–825, Chichester, August 8–12 1994. John Wiley and Sons.
- [8] Sarit Kraus, Eithan Ephrati, and Daniel Lehmann. Negotiation in a non-cooperative environment. *Journal of Experimental and Theoretical Artificial Intelligence*, 3(4):255–282, 1991.
- [9] Martha E. Pollack. The uses of plans. In *Transcript of a talk given on receipt of the Computers and Thought Award at AAAI91*, 1991.
- [10] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: an introduction*. The MIT Press, 1998.
- [11] M. Wooldridge and N. R. Jennings. Towards a theory of cooperative problem solving. In *Proc. Modelling Autonomous Agents in a Multi-Agent World (MAAMAW-94)*, pages 15–26, Odense, Denmark, 1994.
- [12] Michael Wooldridge. *An introduction to Multi-Agent Systems*. Wiley, 2002.